Constrained Profile Retrieval applied to MIPAS Observation Mode

Tilman Steck[1] and Thomas von Clarmann[2]

[1] When this research was performed, T. Steck (Tilman.Steck@jpl.nasa.gov) was with the Deutsches Zentrum für Luft- und Raumfahrt (DLR) e.V., Oberpfaffenhofen, 82234 Weßling, Germany. He is now with the Jet Propulsion Laboratory, California Institute of Technology, M.S. 183-301, 4800 Oak Grove Drive, Pasadena, CA 91109.

[2] Forschungszentrum Karlsruhe, IMK, Postfach 3640, 76344 Karlsruhe, Germany.

Abstract

To investigate the atmosphere of the Earth, and to detect changes in our environment, the Environmental Satellite (ENVISAT) will be launched by the European Space Agency (ESA) in a polar orbit in the mid of the year 2001. One of its payload instruments is the Fourier spectrometer MIPAS (Michelson Interferometer for Passive Atmospheric Sounding), designed to measure the spectral thermal emission of molecules in the atmosphere in a limb viewing mode. The goal of this experiment is to derive vertical profiles of pressure and temperature, as well as of the trace gases O_3, H_2O, CH_4, N_2O, and HNO_3 from

the spectra operationally on a global scale. A major topic in the analysis of the computational methodology to obtain the profiles is how available a priori knowledge can be used, and how this a priori knowledge affects the corresponding results. Retrieval methods were compared and it was shown that the optimal estimation formalism can be used in a highly flexible way for this kind of data analysis. Beyond this, diagnostic tools, such as estimated standard deviation, vertical resolution, or degrees of freedom, have been used to characterize the results. Optimized regularization parameters have been determined and the huge influence of the choice of the regularization and discretization on the result was demonstrated. In particular, we show that the optimal estimation formalism can be used to emulate purely smoothing constraints.

OCIS codes: 000.3860 Mathematical methods in physics, 000.4430 Numerical approximation and analysis, 280.0280 Remote sensing, 300.6340 Spectroscopy, infrared.

Additional Key Words: Retrieval; Regularization.

1 Introduction

Remote sensing of the atmosphere, especially from satellites, allows to determine the global composition of the atmosphere and their temporal changes. For this purpose the European Space Agency (ESA) will launch the European

environmental satellite ENVISAT [1] in the mid of 2001. One of its payload instruments is the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) (cf. [2], [3]). MIPAS is a limb viewing Fourier-Spectrometer measuring the infrared emission of the molecules in the atmosphere. The goal of the MIPAS experiment is to measure vertical profiles of pressure, temperature and mixing ratios of O_3, H_2O, CH_4, N_2O, and HNO_3.

To obtain these target quantities from the measurements, an inverse problem has to be solved. There exist various methods to solve such kind of problems (cf. [4], [5], [6]). In most cases it is set up as a least squares problem, where the different solution formulas are quite similar. In this study different applications of the optimal estimation formalism [7] are assessed. The impact of constraints is shown analytically as well as in numerical experiments under consideration of diagnostic tools such as retrieval error [8], vertical resolution [9] and degrees of freedom.

We proceed as follows: The formal description of the inverse problem and the basic definitions are given in Section 2. An overview of the formalism for solving the least squares problem and the retrieval strategy used is presented in Section 3. Furthermore comparisons between optimal estimation and Tikhonov regularization [10] are made. In Section 4 the retrieval result is analyzed with respect to the influence of regularization. The theoretical considerations are applied to a MIPAS observation mode in Section 5. The predicted retrieval characteristics are illustrated in selected examples. Finally, Section 6 gives a summary and some conclusions of our work.

2 Problem Description and Definitions

The retrieval problem in atmospheric limb sounding is to extract vertical profiles of atmospheric state parameters from a limb sequence of emission spectra of different tangent heights. An exact physical description of the problem requires continuous functions: The spectrum is a continuous function with frequency, while vertical profiles to be retrieved are continuous functions with height. Measurements are always of discrete nature. In this sense most of the inverse problems are formally ill-posed or under-determined [5]. To overcome this, the continuous functions of the vertical profiles are discretely sampled and arranged in a state vector $\mathbf{x} \in \mathbb{R}^n$ of unknown quantities, whereby we will apply in the following the terminology used by Rodgers [5]. This vector contains the finite number of unknown quantities which shall be determined in the retrieval. The width of the grid on which the atmospheric state parameters are represented depends both on practical considerations and on the required altitude resolution.

The m-dimensional measurement vector $\mathbf{y}$ contains all measured radiances of a complete limb scan to be considered for inversion. The measurement is superimposed by instrumental random noise. The correlation between the state vector $\mathbf{x}$, the measurement vector $\mathbf{y}$, the m-dimensional noise vector ϵ, and the radiative transfer model $\mathbf{F}$ is given by

$$\mathbf{y} = \mathbf{F}(\mathbf{x}) + \epsilon. \tag{1}$$

While the forward model maps from the state space (vertical profiles) into the

measurement space (emission spectra), we are interested in the inverse mapping, i.e. an appropriate determination of state vector $\mathbf{x}$ from the measurements. In our case the inverse problem is formally over-determined, i.e. the number of measurements is higher than the number of state parameters ($m > n$). Therefore the inverse problem can be formulated as a least squares problem:

$$\|\mathbf{y} - \mathbf{F}(\mathbf{x})\|^2_{\mathbf{S}_\epsilon^{-1}} = (\mathbf{y} - \mathbf{F}(\mathbf{x}))^T \mathbf{S}_\epsilon^{-1} (\mathbf{y} - \mathbf{F}(\mathbf{x})), \tag{2}$$

where $\mathbf{S}_\epsilon$ is the covariance matrix of the measurement. The nonlinearity of $\mathbf{F}(\mathbf{x})$ often requires Eq. (2) to be solved by Newtonian iteration:

$$\mathbf{x}_{i+1} = \mathbf{x}_i + (\mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} \mathbf{K}_i)^{-1} \mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} (\mathbf{y} - \mathbf{F}(\mathbf{x}_i)), \tag{3}$$

where $\mathbf{K}$ is the Jacobian of $\mathbf{F}$ with respect to $\mathbf{x}$ and the subscript i denotes the iteration index. This solution uses information only from the measurement. Additional knowledge about the solution can be added by inclusion of a constraining term

$$\begin{aligned} &\|\mathbf{y} - \mathbf{F}(\mathbf{x})\|^2_{\mathbf{S}_\epsilon^{-1}} + \mathcal{R}(\mathbf{x}, \mathbf{x}_a, \mathbf{R}) \\ &= (\mathbf{y} - \mathbf{F}(\mathbf{x}))^T \mathbf{S}_\epsilon^{-1} (\mathbf{y} - \mathbf{F}(\mathbf{x})) + (\mathbf{x} - \mathbf{x}_a)^T \mathbf{R} (\mathbf{x} - \mathbf{x}_a), \end{aligned} \tag{4}$$

where $\mathcal{R}$ is the constraining function, $\mathbf{x}_a$ an *a priori* known state vector and $\mathbf{R}$ a regularization matrix. The related iterative solution is given by

$$\begin{aligned} \mathbf{x}_{i+1} \quad = \quad & \mathbf{x}_i + (\mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} \mathbf{K}_i + \mathbf{R})^{-1} \times \\ & [\mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} (\mathbf{y} - \mathbf{F}(\mathbf{x}_i)) - \mathbf{R}(\mathbf{x}_i - \mathbf{x}_a)]. \end{aligned} \tag{5}$$

Depending on the choice of the regularization matrix, different kinds of constraints can be included in the least squares problem. In the following section

two possible choices of the regularization term will examinated in detail and discussed against each other.

3 Retrieval strategy

The iterative solution given by Eq. (5) can be seen as a weighted mean between the information contained in the measurement and the information added by the constraint. Determination of the appropriate weight of the constraint is a critical part in the design of a retrieval strategy.

In the following we will examine two kinds of constraints in more detail: Optimal estimation [7] and Tikhonov regularization [10]. In case of optimal estimation, the difference between the *a priori* profile $\mathbf{x}_a$ and the solution is kept small. The regularization matrix is $\mathbf{R} = \mathbf{S}_a^{-1}$, the inverse of the *a priori* covariance matrix, and the iterative solution is

$$\begin{aligned} \mathbf{x}_{i+1} &= \mathbf{x}_i + (\mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} \mathbf{K}_i + \mathbf{S}_a^{-1})^{-1} \times \\ &\quad [\mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} (\mathbf{y} - \mathbf{F}(\mathbf{x}_i)) - \mathbf{S}_a^{-1} (\mathbf{x}_i - \mathbf{x}_a)]. \end{aligned} \tag{6}$$

Has the inverse of the covariance matrix for the measured profile $\mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{K}$ larger entries than the inverse of the *a priori* covariance matrix $\mathbf{S}_a^{-1}$, the retrieval is dominated by the measurement with its noise. The same considerations can be applied vice versa. If the *a priori* information comes from an independent measurement or if $\mathbf{x}$ is part of a climatological ensemble following Gaussian statistics which is represented by $\mathbf{x}_a$ and $\mathbf{S}_a$, the optimal estimation solution is the most probable solution [7]. However, if there is no independent measurement which

provides $\mathbf{x}_a$ and if $\mathbf{x}$ is not part of the statistical ensemble which is represented by $\mathbf{x}_a$ and $\mathbf{S}_a$, the result is to some degree biased towards the *a priori* profile. A solution constrained in its shape only rather than in its absolute values may be more desirable in order not to miss unexpected atmospheric phenomena. For this purpose, Tikhonov regularization is a powerful tool.

In case of Tikhonov regularization the regularization matrix becomes $\mathbf{R} = \alpha \mathbf{L}^T \mathbf{L}$, where α is a regularization parameter and $\mathbf{L}$ a regularization matrix. This kind of regularization is also known as Twomey [11] or Phillips [12] regularization, and the iterative solution is

$$\begin{aligned} \mathbf{x}_{i+1} &= \mathbf{x}_i + (\mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} \mathbf{K}_i + \alpha \mathbf{L}^T \mathbf{L})^{-1} \times \\ &\quad [\mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} (\mathbf{y} - \mathbf{F}(\mathbf{x}_i)) - \alpha \mathbf{L}^T \mathbf{L} (\mathbf{x}_i - \mathbf{x}_a)]. \end{aligned} \tag{7}$$

The parameter α determines the weight of the regularization and it is also important to chose α appropriate to the problem. One way to fix this parameter is the L-curve method [13].

There are several possibilities to set up the regularization matrix $\mathbf{L}$. The simplest one is the identity matrix, $\mathbf{L} = \mathbf{I}$. This results in a retrieval equivalent to optimal estimation with a diagonal *a priori* covariance matrix $\mathbf{S}_a = \frac{1}{\alpha}\mathbf{I}$. In order to obtain a smooth solution the discrete first derivative operator is useful as smoothing matrix, $\mathbf{L} = \mathbf{L}_1$. The discrete first derivative operator can be derived as follows: the first derivative can be approximated as

$$x'_k \approx \frac{x_k - x_{k-1}}{z_k - z_{k-1}}, \tag{8}$$

where x_k is the state parameter at height z_k. By approximation of the first

derivative by a differential quotient and in case of equidistant altitude grid the operator is defined as

$$\begin{pmatrix} x^{2'} \\ \vdots \\ x^{n'} \end{pmatrix} \approx \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} =: \mathbf{L}_1 \mathbf{x}, \tag{9}$$

where $\mathbf{L}_1 \in \mathbb{R}^{(n-1)\times n}$. This and higher order kind of operator acts just in smoothing and not in biasing the solution.

In the case that two different types of state parameters being jointly fitted which require different approaches of regularization (optimal estimation and smoothing without biasing) it is desirable to have one formalism which supports both types of constraints. In order to gain a deeper understanding of the smoothing effect of the off-diagonal elements of $\mathbf{S}_a$ we try to express $\alpha \mathbf{L}_1^T \mathbf{L}_1$ by a matrix of the format of an inverse covariance matrix.

A covariance matrix consists of diagonal (variances) and off-diagonal (covariances or non-normalized correlations) elements. In the case only the variances of an *a priori* profile are known, covariances are often described by an exponential decay

$$[\mathbf{S}_a]_{ij} = \sqrt{[\mathbf{S}_a]_{ii}[\mathbf{S}_a]_{jj}}\ \exp(-\frac{|z_i - z_j|}{w}),\ \forall\, i, j, \tag{10}$$

where w is a length determining the correlation between the parameters at different heights z_i. In case of equidistant altitude grid and variances of all the same size (assumed for clearer presentation but without affecting the conclu-

sions) Eq. (10) reads

$$[\mathbf{S}_a]_{ij} = d\,\exp(-|i-j|\frac{\Delta}{w}),\ \forall\, i,j. \tag{11}$$

If $\frac{\Delta}{w}$ is chosen very large, the non-diagonal elements of $\mathbf{S}_a$ are very small and the parameters in the vertical profile are only weakly correlated. This corresponds to the case that the *a priori* covariance matrix is diagonal and the corresponding Tikhonov matrix is identity. The case $\frac{\Delta}{w} \ll 1$ is for our application of more interest because the off-diagonal elements are just unessentially smaller than the diagonal elements, and the state parameters are strongly correlated.

The covariance matrix above can be inverted and we obtain

$$\mathbf{S}_a^{-1} = \frac{1}{d} \cdot \frac{1}{1-\exp(-2\frac{\Delta}{w})} \times \tag{12}$$

$$\begin{pmatrix} 1 & -\exp(-\frac{\Delta}{w}) & 0 & \cdots & 0 \\ -\exp(-\frac{\Delta}{w}) & 1+\exp(-2\frac{\Delta}{w}) & -\exp(-\frac{\Delta}{w}) & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \cdots & -exp(-\frac{\Delta}{w}) & 1+\exp(-2\frac{\Delta}{w}) & -\exp(-\frac{\Delta}{w}) \\ 0 & \cdots & 0 & -\exp(-\frac{\Delta}{w}) & 1 \end{pmatrix}.$$

For the case of $\frac{\Delta}{w} \ll 1$ the exponential function in Eq. (11) and (12) can be expanded at zero:

$$\exp(-\frac{\Delta}{w}) \approx 1 - \frac{\Delta}{w}. \tag{13}$$

Putting this approximation into Eq. (12) and neglecting the addends which

imply $\frac{\Delta}{w}$, the inverse of the *a priori* covariance matrix can be expressed as

$$\mathbf{S}_a^{-1} \approx \frac{1}{d}\frac{w}{2\Delta}\begin{pmatrix} 1 & -1 & 0 & \cdots & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 2 & -1 \\ 0 & \cdots & \cdots & 0 & -1 & 1 \end{pmatrix} = \alpha \mathbf{L}_1^T \mathbf{L}_1. \qquad (14)$$

This means that an *a priori* covariance matrix can be set up that it acts approximately in the same way as the discrete first derivative operator $\mathbf{L}_1$. Certainly this equivalence is only approximate, since truncation of the expansion causes a regular matrix $\mathbf{S}_a^{-1}$ being transformed to a singular matrix of the type $\mathbf{L}_1^T\mathbf{L}_1$. Therefore the optimal estimation formalism can be used either in the original sense of weighting the measured and the *a priori* profile, or in the sense of smoothing the solution without considerably shifting it towards the *a priori* profile. Due to the approximate equivalence of both methods, we henceforth restrict our description to optimal estimation terminology.

4 Diagnostics

Following Rodgers [5], several powerful diagnostic tools are available and will be used later-on for assessment of the proposed retrieval approach: covariance matrix of the retrieved solution, averaging kernel matrix, and degrees of freedom. The covariance matrix of the solution includes the estimated error of the state

parameters and their inter-level correlations:

$$\hat{\mathbf{S}} = (\mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{K} + \mathbf{S}_a^{-1})^{-1}. \tag{15}$$

Another important quantity to assess the retrieval is vertical resolution. The vertical resolution depends on the sensitivity of the retrieval to changes in the true profile,

$$\mathbf{A} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}} = \mathbf{D}_y \mathbf{K}, \tag{16}$$

where $\hat{\mathbf{x}}$ is the estimated profile in the linear case and $\mathbf{D}_y$ describes the sensitivity of the retrieval concerning changes in the measurement

$$\mathbf{D}_y = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{y}}.$$

Matrix $\mathbf{A}$ usually is called averaging kernel matrix, and is discussed extensively in the literature (cf. [14], [6]). There exists several ways to define the vertical resolution of the retrieval as a function of $\mathbf{A}$ [9]. We define the vertical resolution as the width at the half maximum of the column of the averaging kernel matrix $\mathbf{A}$.

The averaging kernel matrix for the optimal estimation formalism is calculated as

$$\mathbf{A} = (\mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{K} + \mathbf{S}_a^{-1})^{-1} \mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{K}. \tag{17}$$

If there is no constraining part in the least squares problem (Eq. (2)), the averaging kernel matrix is the identity matrix $\mathbf{I}$. In this case altitude resolution is determined by the grid used for representation of $\mathbf{x}$.

The estimated solution for optimal estimation can also be expressed in terms of the averaging kernel. In the linear case ($\mathbf{y} = \mathbf{K}\mathbf{x} + \epsilon$) the gained equation allows

to separate the true parameters and the *a priori* parameter. The estimated solution is [8]

$$\begin{aligned} \hat{\mathbf{x}} &= \mathbf{x}_a + \mathbf{A}(\mathbf{x} - \mathbf{x}_a) + \mathbf{D}_y\epsilon \\ &= \mathbf{A}\mathbf{x} + (\mathbf{I} - \mathbf{A})\mathbf{x}_a + \mathbf{D}_y\epsilon, \end{aligned} \tag{18}$$

where the averaging kernel $\mathbf{A}$ is the weight of the true state, the matrix $(\mathbf{I} - \mathbf{A})$ the weight of the *a priori* state and $\mathbf{D}_y$ the weight of the noise. Again, in case of no constraints ($\mathbf{A} = \mathbf{I}$) the *a priori* part of Eq. (18) disappears and the result only depends on the true state and the noise.

Because the matrices of the weights $\mathbf{A}$, $(\mathbf{I} - \mathbf{A})$, and $\mathbf{D}_y$ are usually all non-diagonal, it is not trivial to quantify the different parts of the retrieval formula Eq. (18). Rodgers [5] proposed a transformation in a space where the matrices described above become diagonal. For this purpose, the Jacobian $\mathbf{K}$ has to be transformed and a singular value decomposition (SVD) is performed [15]:

$$\tilde{\mathbf{K}} = \mathbf{S}_\epsilon^{-\frac{1}{2}}\mathbf{K}\mathbf{S}_a^{\frac{1}{2}} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T, \tag{19}$$

where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices and $\mathbf{\Lambda}$ is a diagonal matrix with the singular values λ_i of $\tilde{\mathbf{K}}$ as its diagonal elements. Furthermore is $\mathbf{U}$ a $m \times n$ matrix, $\mathbf{V}$ and $\mathbf{\Lambda}$ are $n \times n$ matrices. The state vector and the noise vector are transformed as follows:

$$\begin{aligned} \mathbf{x}' &= \mathbf{V}^T\mathbf{S}_a^{-\frac{1}{2}}\mathbf{x} \\ \epsilon' &= \mathbf{U}^T\mathbf{S}_\epsilon^{-\frac{1}{2}}\epsilon. \end{aligned} \tag{20}$$

With the singular value decomposition of $\tilde{\mathbf{K}}$ (Eq. (19)) and Eq. (20) we receive

for the transformed estimated state parameters (see appendix A)

$$\hat{\mathbf{x}}' = [\mathbf{I} + \mathbf{\Lambda}^2]^{-1}\mathbf{\Lambda}^2\mathbf{x}' + [\mathbf{I} + \mathbf{\Lambda}^2]^{-1}\mathbf{x}'_a + [\mathbf{I} + \mathbf{\Lambda}^2]^{-1}\mathbf{\Lambda}\epsilon'. \tag{21}$$

The matrices in the different addends are diagonal now. Therefore the equation above can be written component wise

$$\hat{x}'_i = \frac{\lambda_i^2}{1+\lambda_i^2}x'_i + \frac{1}{1+\lambda_i^2}x'_{a_i} + \frac{\lambda_i}{1+\lambda_i^2}\epsilon'_i. \tag{22}$$

It can be seen from Eq. (22), that the transformed state parameter $\hat{x}'_i$ is dominated by the true profile if the corresponding singular value is much larger than 1 ($\lambda_i \gg 1$). For cases with $\lambda_i \ll 1$ the corresponding state parameter is determined by the *a priori* knowledge. For singular values $\lambda_i \approx 1$, the estimated state parameter is influenced by the true state, the *a priori* state, and the noise with the same weight. The only disadvantage of this kind of representation is that the transformed quantities lose their simple physical meaning.

Eq. (22) leads to the definition of the degrees of freedom for signal [5]:

$$\text{dfs} = \sum_i \frac{\lambda_i^2}{1+\lambda_i^2}. \tag{23}$$

The singular values with $\lambda_i \ll 1$ do not contribute to the degrees of freedom because they will be determined by the *a priori* knowledge. Therefore the number of degrees of freedom is not fixed with the dimension of the state space n but rather with the quantity computed in Eq. (23).

The diagnostic tools where the singular value decomposition of the transformed Jacobian is used, is for special application to the optimal estimation formalism. The singularity of $\alpha\mathbf{L}_1^T\mathbf{L}_1$ makes it necessary to use different analysis which is described in Hansen [13].

5 Application to a MIPAS Observation Mode

Retrieval methods and diagnostic tools discussed in the previous Sections will be applied to simulated methane and ozone retrievals typical for MIPAS on ENVISAT.

5.1 General assumptions

The forward modeling is performed with the Karlsruhe Optimized and Precise Radiative transfer Algorithm (KOPRA). Details and characteristics of KOPRA can be found in [16], [17], and [18]. The forward model KOPRA is used in the retrieval part of the MIPAS Level 2 off-line processor (cf. [19], [20], and [21]). Synthetic measurements have been generated by calculating a limb sequence of reference spectra by KOPRA, which then were superimposed by synthetically generated noise of realistic amplitude and covariance. Other instrumental characteristics like field-of-view (FOV) and spectral resolution were also considered (see Tab. 1). For the additional apodization of the spectra the strong version of the Norton-Beer apodization was used. Non-local thermodynamic equilibrium was not considered even though the forward model KOPRA is able to model it. In contrast to most other forward models, the Earth's ellipticity was taken into account. The limb measurements were simulated at 16 tangent heights from 53 km down to 8 km in 3 km steps. The actual nominal observation mode was not used here because the existing ESA studies rely on the measurement mode described above and the possibility of comparisons were quite limited.

From the entity of measurements a spectral subset, so-called microwindows [22],

has been selected for analysis. The occupation matrices for methane and ozone, shown in Tables 2 and 3, contain the information which microwindow is used at which tangent height. The dimension of the measurement vector for methane is $m = 3958$ and $m = 2085$ for ozone, respectively. Furthermore we used pressure, temperature and constituent profiles as provided by [23] to generate the synthetic measurements as well as for initial guess and *a priori* profiles. These profiles will be called "standard profiles" in the following.

The retrieval is performed on an fine altitude grid, i.e. a grid from 8 km to 53 km with a grid width of 1 km. Therefore the dimension of the state vector is $n = 46$. The fine grid allows to represent small scale vertical structures in the vertical profiles; this large number of unknowns to be retrieved from measurements at only 16 tangent altitudes leads to an ill-posed problem. Therefore regularization is required. We apply the optimal estimation formalism in the smoothing sense (Eq. (14)). This corresponds to large off-diagonal values of $\mathbf{S}_a$ with the intention to smooth the solution rather than biasing it.

5.2 Application to methane

Methane has strong lines in the considered infrared region. The retrieval of the vertical profile should lead to results with a relative error smaller than 10 % in most altitude regions. The power of the developed method, optimal estimation without biasing, is demonstrated in this example. The reference profile for methane to generate the synthetic data is a scaled version of the standard profile by a factor of 0.8.

Fig. 1 shows the result of the methane retrieval for an *a priori* covariance matrix $\mathbf{S}_a$ which is constant with height. In this example the diagonal is $[\mathbf{S}_a]_{ii} = d = 10.0$ ppmv^2 $\forall\, i$, and the correlation parameter is $w = 2000.0$ km. The volume mixing ratio in units of parts per million by volume (ppmv) is depicted versus altitude. The dotted line represents the *a priori* and the initial guess profile, the starting point of the iteration. The dashed line shows the scaled methane reference profile which had been used to create the simulated measurements. The solid line represents the retrieved profile. The reference and the retrieved profile show a good agreement which leads to an absolute root-mean-square (abs. rms) error of 0.0306 ppmv. This quantity is computed as follows:

$$\mathrm{rms}_{\mathrm{abs}} = \sqrt{\sum_{i=1}^{n}(x_i - x_{\mathrm{ref},i})^2/n}. \tag{24}$$

In Fig. 2 the absolute deviation between the reference profile and the solution is illustrated as a function of altitude. The solid line shows the empirical deviation, which depends on the actual noise ϵ the simulated measurement is superimposed with. It is nearly at all heights less than 0.05 ppmv. The dashed curve shows the estimated absolute deviation as evaluated by Eq. (15). The displayed values are the square roots of the diagonal elements of the solution covariance matrix $\hat{\mathbf{S}}$. The estimated deviation envelopes the real deviation quite well. Furthermore it is obvious, that the estimated deviation is nearly constant with height. This corresponds with the height constant *a priori* covariance matrix $\mathbf{S}_a$ which dominates the shape of the estimated deviation.

Also the altitude distribution of the relative deviation (Fig. 3) is instructive. It

is computed by

$$\mathrm{rms}_{\mathrm{rel}} = \sqrt{\sum_{i=1}^{n} \left(\frac{x_i - x_{\mathrm{ref}_i}}{x_{\mathrm{ref}_i}} \right)^2 / n}. \tag{25}$$

and the related relative root-mean-square is 4.69 %. The real relative deviation (solid curve in Fig. 3) is smaller than 10 % at all heights. The relative estimated standard deviation, which is the the square-root of the diagonal of Eq. (15) divided by the reference profile, increases continuously with height. The reason of this behavior is the decreasing methane amount with height (Fig. 1) and the height-independent absolute deviation (Fig. 2).

In order to evaluate the influence of the regularization on this result, we will use the methods derived in Section 4. In Fig. 4 the averaging kernel is displayed in a scaled form, where the bright regions correspond to large values and dark regions to small ones. The matrix is dominated by the diagonal, that means that the retrieved state parameters reflect the true profile quite well. But the smoothing influence of the *a priori* covariance matrix produces also off-diagonal elements. The single columns of the averaging kernel are shown in Fig. 5. The width of the columns of **A** (vertical resolution) is displayed in Fig. 6. The mean vertical resolution which is driven by the chosen regularization is 3.31 km and exceeds slightly the distance of two consecutive tangent heights (3 km). However, the chosen constraint lets enough variability to detect small-scaled structures.

We will complete the diagnostic of the retrieval with the singular value decomposition of the transformed Jacobian (Eq. (19)). The singular values are displayed in order of size in Fig. 7. The 16th value lies directly above the dashed line corresponding to the value 1. Therefore most of the singular values are smaller

than one and the degrees of freedom are calculated with Eq. (23) to 15.9. This means that only about 16 parameters of the state vector is determined by the measurement and the remaining 30 are determined by the constraint. This result is not surprising, rather the basis for the choice of the constraint. Because we used an observation scenario with 16 tangent heights, the vertical resolution of MIPAS on ENVISAT is 3 km, and the fact that the spectral resolution in this case is not high enough to gain additional information of values between the tangent heights, we do not expect more degrees of freedom than the number of tangent heights. This conclusion is supported by the characteristics of Fig. 5, where it is remarkable that always three columns are peaked at nearly the same height. That means the main information in the measured spectra of MIPAS ENVISAT originates at the tangent heights (distance between two height is 3 km) and the two retrieved parameters above the tangent heights correlates with the actual values.

This fact is also the explanation for the fluctuating of the estimated deviations (Figures 2 and 3) and the values of the vertical resolution (Fig. 6). The worse estimated deviations and the worse vertical resolution correspond with the retrieval parameter at an altitude of 2 km above a tangent point. Therefore at these points is least information about the target species available. The retrieval parameters at an altitude of 1 km above a tangent height contain nearly as much information as at the tangent point itself (see Figures 2 and 3).

The proportion of the measurement and the *a priori* can also be seen in Fig. 8. As expected the two curves intersect between the 16th and 17th singular value

λ_i. The constraint acts like a truncated SVD, where the expansion is not "hard" truncated at a single value but "smooth" in a specific range.

To sum up the diagnostic of the retrieval with height constant regularization for methane, it can be seen that both the absolute deviation (0.0306 ppmv) and the relative deviation (4.69 %) are very small. The example demonstrates that the applied method leads to stable solution which can be analyzed in detail with the available diagnostic tools.

5.3 Height-dependent and -independent constraints applied to ozone

The regarded infrared spectral range has also strong ozone lines. For reference, an ozone hole profile was used in order to assess the capability of the algorithm to retrieve highly structured vertical profiles from MIPAS measurements.

The result of the ozone retrieval for an *a priori* covariance matrix $\mathbf{S}_a$ which is constant with height is shown in Fig. 9. The diagonal is set to $[\mathbf{S}_a]_{ii} = d = 100.0$ ppmv^2 $\forall\, i$, and the correlation parameter is $w = 5000.0$ km. The dotted line represents again the *a priori* and the initial guess profile. Note that the *a priori* profile has a significant different shape in comparison to the reference profile (dashed line), which is the profile under ozone hole conditions. It is characterized by a reduced amount of ozone between 15 km and 40 km. The visual good agreement between the reference and the retrieved profile (solid line) is manifested in a small absolute root-mean-square (abs. rms) error of 0.111 ppmv.

In Fig. 10 the absolute deviation between the reference profile and the solution is highlighted. It exceeds the value of 0.2 ppmv only at altitudes of 20 and 23 km. The estimated deviation (dashed curve) envelopes the real deviation quite well. As mentioned in the methane case, it is obvious again that the estimated deviation is nearly constant with height and that the height constant *a priori* covariance matrix $\mathbf{S}_a$ dominates the shape of the estimated deviation. The altitude distribution of the relative deviation (Fig. 11) leads to a result (rel.rms.: 29.2 %) which is not desired. The real relative deviation (solid curve in Fig. 11) is smaller than 10 % above 20 km. At lower altitudes where the absolute values of the ozone profile become quite small, the relative deviation exceeds 50 % both in terms of estimated standard deviation and deviation of the actual retrieval. This suggests that the chosen approach of regularization constant with altitude may be non-satisfactory at least for retrieval problems with emphasis on lower stratosphere – upper troposphere.

For a further diagnostic of this result we will focus on the vertical resolution which is displayed in Fig. 12. The mean vertical resolution which is driven by the chosen regularization is 3.42 km and exceeds the distance of two consecutive tangent heights (3 km). Especially in the range between 20 km and 50 km the vertical resolution is 3.5 km and worse. The smoothing by the retrieval in this region seems to be too strong. In contrast, below 15 km the smoothing is quite small. Here the vertical resolution is pushed towards values smaller than 2 km, at the cost of large random errors (Fig. 11). This clearly is a result of under-regularization at altitudes below 15 km. A stronger regularization in this height

range should lead to smaller relative deviations.

Therefore we will choose a height dependent *a priori* covariance matrix on purpose to receive a smaller relative deviation in the lower part of the ozone profile. In contrast to the regularization above, we now assess the behavior of the retrieval where an altitude-dependent regularization matrix is applied. The variances are set proportional to the squared volume mixing ratios of the related altitude (see Tab. 4), and the decay towards off-diagonal elements according to Eq. (13) involves a correlation parameter of $w = 2.0 \times 10^6$ km.

Again the dotted line in Fig. 13 represents the *a priori* and the initial guess profile and the dashed line the ozone-hole reference profile, like in the previous example. The solid line shows the retrieved ozone profile for the height dependent regularization which matches the reference profile quit well. But the absolute rms error is 0.176 ppmv and is more than 50 % larger than in the height constant case (see Fig. 9). As it can be seen in Fig. 14 the real absolute deviation (solid line) is often larger than 0.2 ppmv in the region between 20 km and 40 km. Nearly all values of the real absolute deviation lie inside the estimated standard deviation. This means that the estimated error is considerable larger than the real error. As expected, the height-dependent constraint does not improve the absolute error, while the relative rms profile error is reduced considerably (Fig. 15) to 13.6 %. Especially in the altitude region below 20 km the real relative deviation (solid line in Fig. 15) is much smaller than in the previous case. The goal of reducing the relative error was successfully achieved at the cost of increased absolute error. In the following we will explain this

behavior.

For a regularization matrix $\mathbf{S}_a$ constant with altitude, the regularization term in optimal estimation

$$(\mathbf{x} - \mathbf{x}_a)^T \mathbf{S}_a^{-1} (\mathbf{x} - \mathbf{x}_a) \tag{26}$$

becomes a quantity which stresses the absolute difference between the estimated and the *a priori* profile in a height-independent way. In the case of the ozone profile, where the absolute vmr values can differ considerably between 8 km and 53 km, the absolute constraint will be the same for all heights. This means that 0.1 ppmv error at 8 km is considered equal a 0.1 ppmv error at 35 km. For this reason the estimated absolute deviation is nearly height constant (Fig. 14) and the absolute rms is very small. Therefore a height constant constraint minimizes the absolute deviation. In contrast to that a well-chosen height dependent *a priori* covariance matrix $\mathbf{S}_a$ will minimize the relative deviation. The differences in Eq. (26) are weighted with the scaled absolute values of the *a priori* profile. Therefore the percentual deviation at 8 km has the same weight as at 53 km. The weight of the absolute deviation will not be so strong which cause large absolute errors in the region between 20 and 40 km (see Fig. 13 and Fig. 14).

The choice of the *a priori* covariance matrix will substantially influence the retrieval result. Especially a height constant or a height dependent constraint will minimize the absolute or the relative deviation. So the choice of the constraint depends on the requirement on the result. We also performed retrieval calculations where the *a priori* covariance matrix is only less height dependent. As

expected, the absolute and the relative errors lie between the errors presented here.

The height-dependent regularization leads to a vertical resolution which is nearly independent from height (see Fig. 16). There are only some altitudes where the resolution is worse than 3.5 km. The mean vertical resolution is calculated to 3.12 km and is slightly smaller than in the height constant case (3.42 km).

6 Summary and Conclusions

The inverse problem of retrieving a vertical atmospheric constituent profile from spectroscopic limb radiance measurements was formulated as a least squares problem. A constraining term of a quadratical form was added to the iterative solution formula in order to obtain a stable solution on a fine altitude grid. The optimal estimation method and the Tikhonov regularization have been assessed in more detail. The only difference between these two kinds of retrieval methods is the regularization matrix $\mathbf{R}$, which is $\mathbf{S}_a^{-1}$ in the optimal estimation case and $\alpha \mathbf{L}^T \mathbf{L}$ for Tikhonov. It was shown that for a special choice of the *a priori* covariance matrix $\mathbf{S}_a$ the discrete first derivative operator $\mathbf{L}_1$ can be simulated. For this kind of *a priori* covariance matrix the solution is not biased but just smoothed.

To characterize the gained retrieval result accurately, several diagnostic tools were used. Which retrieval errors are expected because of noisy measurements tell us the estimated absolute and relative deviation. The composition of the

state parameter at one height is analyzed by the averaging kernel and the corresponding vertical resolution. And finally a reliable quantity of the degrees of freedom for signal is determined by the singular value decomposition of the transformed Jacobian. It could be shown that all these quantities give a complete analysis of the retrieval concerning noisy measurements and regularization in the retrieval formula.

The derived retrieval formula and the diagnostic tools were applied to typical MIPAS ENVISAT methane and ozone retrievals. For the case of methane, the power of the developed retrieval technique could be demonstrated by means of the available diagnostic tools which allow a detailed analysis of the retrieved profile. For ozone, a reference profile was chosen, which is typical for ozone-hole conditions. The calculations were performed for a height constant and a height dependent *a priori* covariance matrix $\mathbf{S}_a$. It could be shown the former will minimize the absolute and the latter the relative rms error. Depending on the requirement on the solution, the constraining term can be adapted.

Appendix A: Transformation of the Linear Retrieval Equation

The retrieval equation for optimal estimation in the linear case (Eq. (18))

$$\hat{\mathbf{x}} = \mathbf{A}\mathbf{x} + (\mathbf{I} - \mathbf{A})\mathbf{x}_a + \mathbf{D}_y\epsilon$$

depending on the averaging kernel

$$\mathbf{A} = [\mathbf{K}^T\mathbf{S}_\epsilon^{-1}\mathbf{K} + \mathbf{S}_a^{-1}]^{-1}\mathbf{K}^T\mathbf{S}_\epsilon^{-1}\mathbf{K}$$

and the sensitivity functions

$$\mathbf{D}_y = [\mathbf{K}^T\mathbf{S}_\epsilon^{-1}\mathbf{K} + \mathbf{S}_a^{-1}]^{-1}\mathbf{K}^T\mathbf{S}_\epsilon^{-1}$$

shall be transformed that the matrices $\mathbf{A}$ and $\mathbf{D}_y$ become diagonal. The following transformations are necessary to do this:

$$\tilde{\mathbf{K}} = \mathbf{S}_\epsilon^{-\frac{1}{2}}\mathbf{K}\mathbf{S}_a^{\frac{1}{2}} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T \tag{27}$$

$$\mathbf{x}' = \mathbf{V}^T\mathbf{S}_a^{-\frac{1}{2}}\mathbf{x} \tag{28}$$

$$\epsilon' = \mathbf{U}^T\mathbf{S}_\epsilon^{-\frac{1}{2}}\epsilon. \tag{29}$$

With $\tilde{\mathbf{K}}^T\tilde{\mathbf{K}} = \mathbf{S}_a^{\frac{1}{2}}\mathbf{K}^T\mathbf{S}_\epsilon^{-1}\mathbf{K}\mathbf{S}_a^{\frac{1}{2}}$, the properties of the SVD, and (27) the averaging kernel becomes

$$\begin{aligned}
\mathbf{A} &= [\mathbf{S}_a^{-\frac{1}{2}}\tilde{\mathbf{K}}^T\tilde{\mathbf{K}}\mathbf{S}_a^{-\frac{1}{2}} + \mathbf{S}_a^{-1}]^{-1}\mathbf{S}_a^{-\frac{1}{2}}\tilde{\mathbf{K}}^T\tilde{\mathbf{K}}\mathbf{S}_a^{-\frac{1}{2}} \\
&= \mathbf{S}_a^{\frac{1}{2}}[\tilde{\mathbf{K}}^T\tilde{\mathbf{K}} + \mathbf{I}]^{-1}\tilde{\mathbf{K}}^T\tilde{\mathbf{K}}\mathbf{S}_a^{-\frac{1}{2}} \\
&= \mathbf{S}_a^{\frac{1}{2}}[\mathbf{V}\mathbf{\Lambda}^2\mathbf{V}^T + \mathbf{I}]^{-1}\mathbf{V}\mathbf{\Lambda}^2\mathbf{V}^T\mathbf{S}_a^{-\frac{1}{2}} \\
&= \mathbf{S}_a^{\frac{1}{2}}\mathbf{V}[\mathbf{\Lambda}^2 + \mathbf{I}]^{-1}\mathbf{\Lambda}^2\mathbf{V}^T\mathbf{S}_a^{-\frac{1}{2}}
\end{aligned} \tag{30}$$

and the sensitivity functions are

$$\begin{aligned}
\mathbf{D}_y &= [\mathbf{S}_a^{-\frac{1}{2}}\tilde{\mathbf{K}}^T\tilde{\mathbf{K}}\mathbf{S}_a^{-\frac{1}{2}} + \mathbf{S}_a^{-1}]^{-1}\mathbf{S}_a^{-\frac{1}{2}}\tilde{\mathbf{K}}^T\mathbf{S}_\epsilon^{-\frac{1}{2}} \\
&= \mathbf{S}_a^{\frac{1}{2}}[\tilde{\mathbf{K}}^T\tilde{\mathbf{K}} + \mathbf{I}]^{-1}\tilde{\mathbf{K}}^T\mathbf{S}_\epsilon^{-\frac{1}{2}} \\
&= \mathbf{S}_a^{\frac{1}{2}}\mathbf{V}[\mathbf{\Lambda}^2 + I]^{-1}\mathbf{\Lambda}\mathbf{U}^T\mathbf{S}_\epsilon^{-\frac{1}{2}}.
\end{aligned} \tag{31}$$

In multiplicating Eq. (18) with $\mathbf{V}^T\mathbf{S}_a^{-\frac{1}{2}}$ and putting (27) - (31) in it we get for the transformed retrieval equation in the linear case

$$\begin{aligned}\hat{\mathbf{x}}' &= [\mathbf{\Lambda}^2 + I]^{-1}\mathbf{\Lambda}^2\mathbf{x}' + (I - [\mathbf{\Lambda}^2 + I]^{-1}\mathbf{\Lambda}^2)\mathbf{x}'_a + [\mathbf{\Lambda}^2 + I]^{-1}\mathbf{\Lambda}\epsilon' \\ &= [\mathbf{\Lambda}^2 + I]^{-1}\mathbf{\Lambda}^2\mathbf{x}' + [\mathbf{\Lambda}^2 + I]^{-1}\mathbf{x}'_a + [\mathbf{\Lambda}^2 + I]^{-1}\mathbf{\Lambda}\epsilon'. \qquad (32)\end{aligned}$$

Acknowledgments. The authors thank Achim Friedle, Klaus Ressel, and Manfred Göbel for their work on the processing software and for helpful discussions and comments.

References

[1] J. Louet. ENVISAT mission & ground segment overview. In *Proceedings of the European Symposium on Atmospheric Measurements from Space, ESAMS'99, 18-22 January 1999, Noordwijk*, volume 1, pages 17–23. European Space Agency, ESTEC, Noordwijk, The Netherlands, 1999.

[2] H. Fischer. Remote sensing of atmospheric trace constituents using Fourier transform spectrometry. *Ber. Bunsenges. Phys. Chem.*, 96(3):306–314, 1992.

[3] H. Fischer. MIPAS mission objectives. In *Proceedings of the European Symposium on Atmospheric Measurements from Space, ESAMS'99, 18-22 January 1999, Noordwijk*, volume 1, page 27. European Space Agency, ESTEC, Noordwijk, The Netherlands, 1999.

[4] R. T. Cox. Probability, frequency and reasonable expectation. *American Journal of Physics*, 14(1):1–13, 1946.

[5] C. D. Rodgers. Inverse methods for atmospheric sounding: Theory and practise. Draft, http://www.atm.ox.ac.uk/user/rodgers/, February 1998.

[6] W. Menke. *Geophysical Data Analysis: Discrete Inverse Theory.* Academic Press, 1984.

[7] C. D. Rodgers. Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation. *Reviews of Geophysics and Space Physics*, 14(4):609–624, 1976.

[8] C. D. Rodgers. Characterization and error analysis of profiles retrieved from remote sounding measurements. *J. Geophys. Res.*, 95(D5):5587–5595, 1990.

[9] A. Dudhia and B. M. Dinelli. Optimisation of the Atmospheric Vertical Grid. Technical report, ESA, 1997. Draft Final Report.

[10] A. Tikhonov. On the solution of incorrectly stated problems and a method of regularization. *Dokl. Acad. Nauk SSSR*, 151:501, 1963.

[11] S. Twomey. On the numerical solution of Fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature. *J. Assoc. Comput. Math.*, 10:97–101, 1963.

[12] D. Phillips. A technique for the numerical solution of certain integral equations of the first kind. *J. Ass. Comput. Math.*, 9:84–97, 1962.

[13] P. C. Hansen. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review*, 34:561–580, 1992.

[14] G. E. Backus and J. F. Gilbert. The resolving power of gross earth data. *Geophys. J. R. Astron. Soc.*, 16:169–205, 1968.

[15] G. H. Golub and W. Kahan. Calculating the singular values and pseudoinverse of a matrix. *SIAM J. Numer. Anal. Ser. B, 2*, pages 205–224, 1965.

[16] M. Höpfner, G. P. Stiller, M. Kuntz, T. v. Clarmann, G. Echle, B. Funke, N. Glatthor, F. Hase, H. Kemnitzer, and S. Zorn. The Karlsruhe optimized

and precise radiative transfer algorithm. Part II: Interface to retrieval applications. In T. Ogawa J. Wang, B. Wu and Z. Guan, editors, *Optical Remote Sensing of the Atmosphere and Clouds, Beijing, China, 15-17 September 1998*, volume 3501 of *Proc. of SPIE*, pages 186–195, 1998.

[17] M. Kuntz, M. Höpfner, G. P. Stiller, T. v. Clarmann, G. Echle, B. Funke, N. Glatthor, F. Hase, H. Kemnitzer, and S. Zorn. The Karlsruhe optimized and precise radiative transfer algorithm. Part III: ADDLIN and TRANSF algorithms for modeling spectral transmittance and radiance. In T. Ogawa J. Wang, B. Wu and Z. Guan, editors, *Optical Remote Sensing of the Atmosphere and Clouds, Beijing, China, 15-17 September 1998*, volume 3501 of *Proc. of SPIE*, pages 247–256, 1998.

[18] G. P. Stiller, M. Höpfner, M. Kuntz, T. v. Clarmann, G. Echle, H. Fischer, B. Funke, N. Glatthor, F. Hase, H. Kemnitzer, and S. Zorn. The Karlsruhe optimized and precise radiative transfer algorithm. Part I: requirements, justification, and model error estimation. In T. Ogawa J. Wang, B. Wu and Z. Guan, editors, *Optical Remote Sensing of the Atmosphere and Clouds, Beijing, China, 15-17 September 1998*, volume 3501 of *Proc. of SPIE*, pages 257–268, 1998.

[19] T. v. Clarmann, G. Stiller, A. Friedle, K. Ressel, and T. Steck. The MIPAS level-2 off-line processor: Requirements and concepts. In *Proceedings of the European Symposium on Atmospheric Measurements from Space,*

ESAMS'99, 18-22 January 1999, Noordwijk, volume 2, pages 529–532. European Space Agency, ESTEC, Noordwijk, The Netherlands, 1999.

[20] A. Friedle, M. Göbel, S. Hilgers, H. Kemnitzer, K. Ressel, G. Schwarz, S. Slijkhuis, T. Steck, T. v. Clarmann, G. Echle, and M. Höpfner. The MIPAS level-2 off-line processor: Realization and test results. In *Proceedings of the European Symposium on Atmospheric Measurements from Space, ESAMS'99, 18-22 January 1999, Noordwijk*, volume 2, pages 509–512. European Space Agency, ESTEC, Noordwijk, The Netherlands, 1999.

[21] K. Ressel. Solving the inverse retrieval problem: A conceptional approach. In *Proceedings of the European Symposium on Atmospheric Measurements from Space, ESAMS'99, 18-22 January 1999, Noordwijk*, volume 2, pages 503–508. European Space Agency, ESTEC, Noordwijk, The Netherlands, 1999.

[22] T. v. Clarmann and G. Echle. Selection of optimized microwindows for atmospheric spectroscopy. *Applied Optics*, 37(33):7661–7669, 1998.

[23] S. A. Clough, F. X. Kneizys, E. P. Shettle, and G. P. Anderson. Atmospheric radiance and transmittance: FASCOD2. In *Proceedings of the Sixth Conference on Atmospheric Radiation*, pages 141–146. Am. Meteorol. Soc., Williamsburg, Va., 1986.

Table 1: Geometrical and spectral characteristics of MIPAS on ENVISAT for the considered observation mode.

Vertical scan range	between 8 and 53 km at tangent height
Vertical resolution	$\approx$ 3 km (width of the FOV)
Spectral range	685 cm^{-1} - 2410 cm^{-1}
	or 4.15 μm - 14.6 μm
Spectral resolution	= 0.035 cm^{-1}
NESR	between 50 nW/(cm^2 sr cm^{-1}) at 685 cm^{-1}
	and 4.2 nW/(cm^2 sr cm^{-1}) at 2410 cm^{-1}

Table 2: Microwindows of CH_4 and the corresponding occupation matrix. The entry 'T' means that the microwindow is used at this tangent height and 'F' means that the microwindow is neglected.

Microwindows	Occupation matrix (tangent heights [km])															
[cm^{-1}]	8	11	14	17	20	23	26	29	32	35	38	41	44	47	50	53
1219.500–1220.000	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
1228.450–1229.050	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
1229.900–1230.400	F	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
1271.300–1271.700	F	F	T	T	T	T	T	T	T	T	T	T	T	T	T	T
1303.300–1303.800	F	F	F	T	T	T	T	T	T	T	T	T	T	T	T	T
1305.550–1306.150	F	F	F	F	T	T	T	T	T	T	T	T	T	T	T	T
1326.900–1327.500	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
1331.975–1332.825	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
1355.000–1356.750	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T

Table 3: Microwindows of O_3 and the corresponding occupation matrix. The entry 'T' means that the microwindow is used at this tangent height and 'F' means that the microwindow is neglected.

Microwindows	Occupation matrix (tangent heights [km])															
[cm^{-1}]	8	11	14	17	20	23	26	29	32	35	38	41	44	47	50	53
722.675– 722.900	F	F	F	F	T	T	T	F	F	F	F	F	F	F	F	T
763.500– 764.650	T	T	T	T	F	T	T	F	F	F	F	F	F	F	F	F
768.400– 768.800	F	T	T	T	T	T	T	F	F	F	F	F	F	F	F	F
777.450– 778.525	T	T	T	F	F	F	F	F	F	F	F	F	F	F	F	F
1031.600–1033.150	F	F	F	F	T	T	T	T	T	T	T	T	T	T	T	T
1052.600–1053.800	T	F	F	F	F	F	T	T	T	F	F	F	F	F	F	T
2120.000–2120.800	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T

Table 4: Height dependent diagonal elements of the *a priori* covariance matrix $\mathbf{S}_a$ for ozone.

Altitude [km]	Variance [ppmv^2]	Altitude [km]	Variance [ppmv^2]
8.0	6.14186×10^1	31.0	8.17027×10^5
9.0	1.45038×10^2	32.0	8.96774×10^5
10.0	2.97483×10^2	33.0	9.61802×10^5
11.0	7.96904×10^2	34.0	1.01022×10^6
12.0	1.65293×10^3	35.0	1.05982×10^6
13.0	2.55242×10^3	36.0	1.05582×10^6
14.0	4.36585×10^3	37.0	1.05183×10^6
15.0	7.30176×10^3	38.0	1.02309×10^6
16.0	1.30639×10^4	39.0	9.70634×10^5
17.0	2.43128×10^4	40.0	9.19558×10^5
18.0	4.34597×10^4	41.0	8.12047×10^5
19.0	7.11091×10^4	42.0	7.11219×10^5
20.0	1.14772×10^5	43.0	6.23278×10^5
21.0	1.58214×10^5	44.0	5.46953×10^5
22.0	2.29512×10^5	45.0	4.75611×10^5
23.0	2.99770×10^5	46.0	3.95917×10^5
24.0	3.69430×10^5	47.0	3.23526×10^5
25.0	4.51995×10^5	48.0	2.62460×10^5
26.0	5.01687×10^5	49.0	2.11383×10^5
27.0	5.53970×10^5	50.0	1.65828×10^5
28.0	6.11512×10^5	51.0	1.39178×10^5
29.0	6.74699×10^5	52.0	1.14861×10^5
30.0	7.40992×10^5	53.0	9.28772×10^4

List of Figures

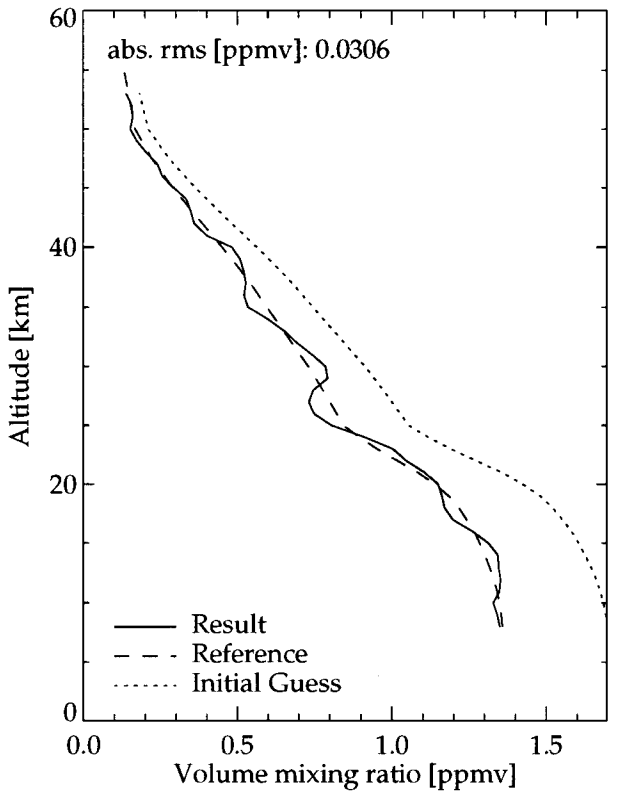

Figure 1: Result of the optimal estimation retrieval in case of an scaled methane reference profile and a regularization constant with height.

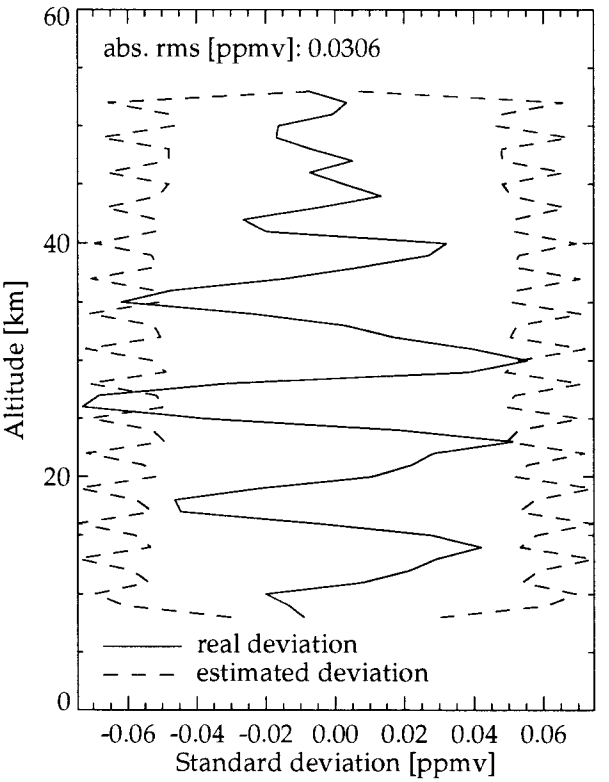

Figure 2: Real and estimated absolute deviation.

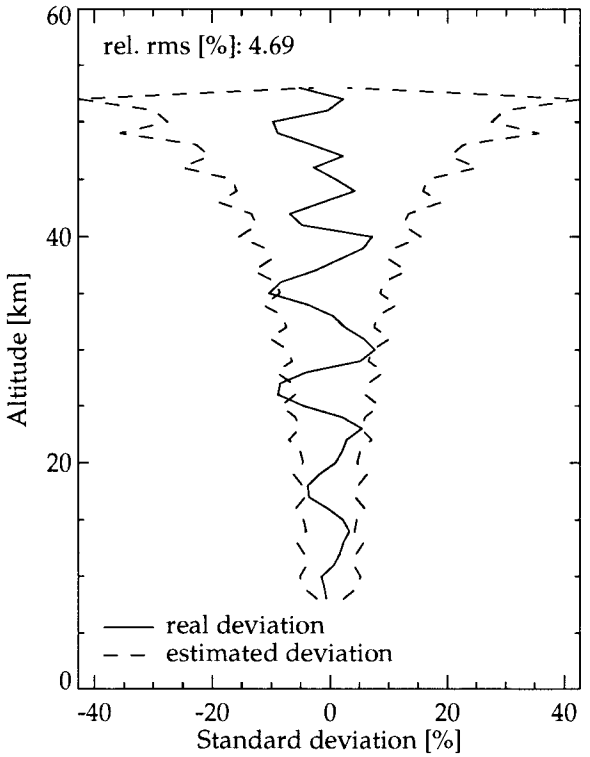

Figure 3: Real and estimated relative deviation.

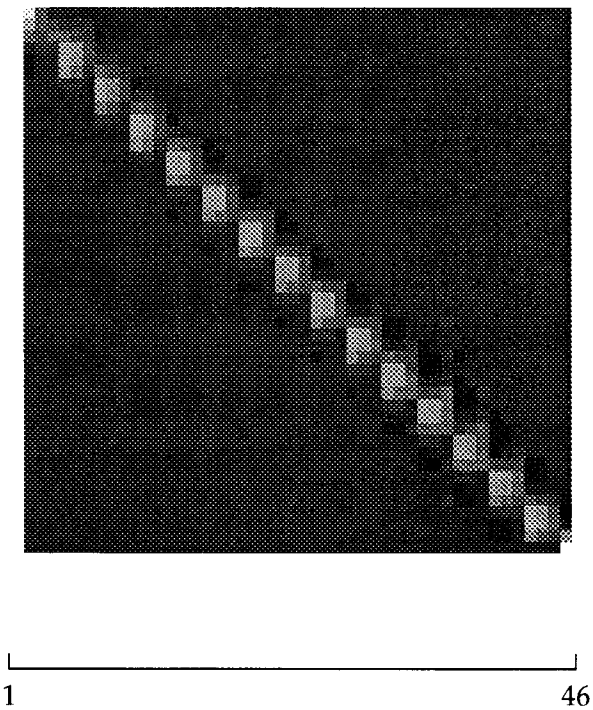

Figure 4: Scaled averaging kernel. The large dark area corresponds to zero-level.

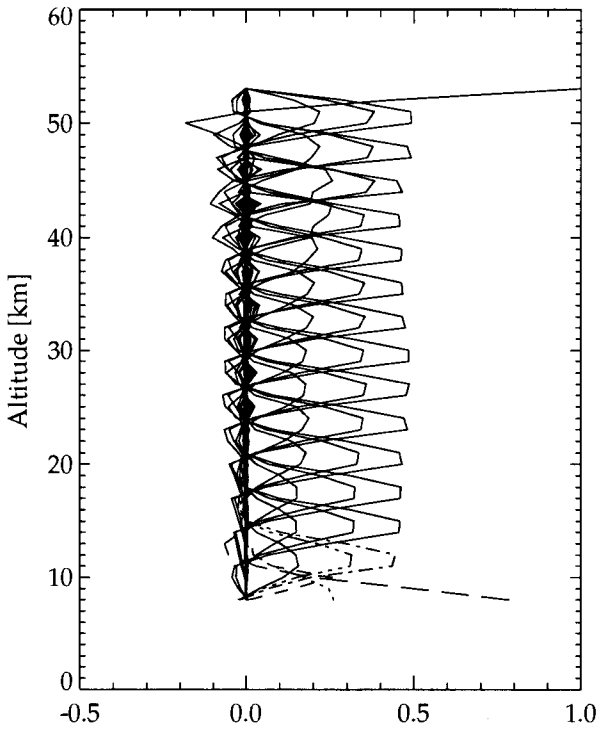

Figure 5: Averaging kernel column-wise.

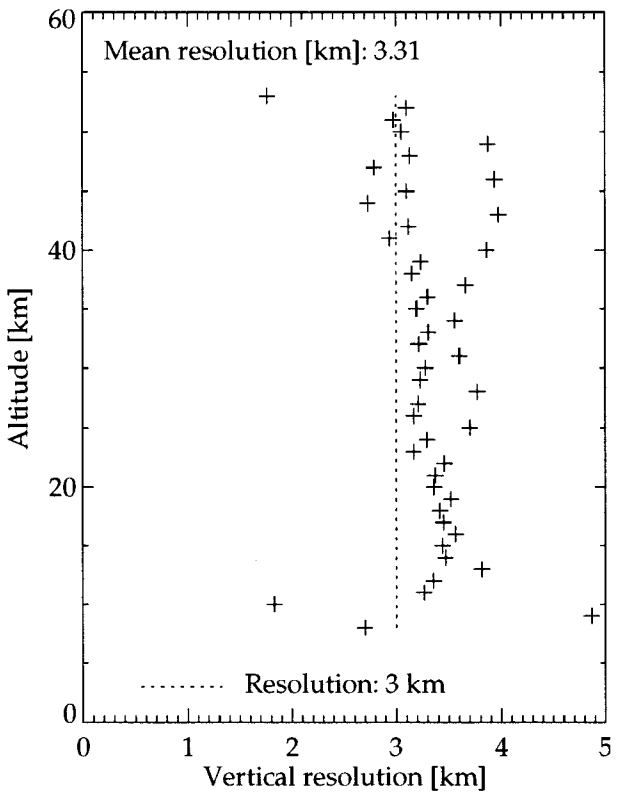

Figure 6: Vertical resolution computed via the widths of the columns of the averaging kernel.

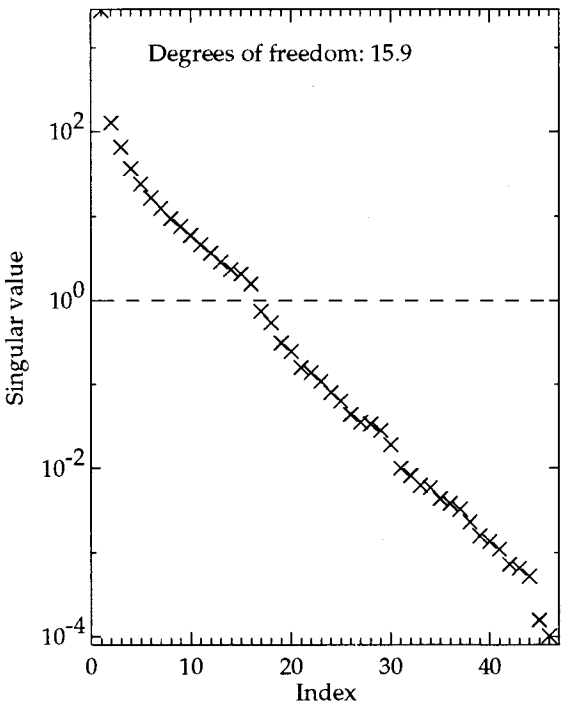

Figure 7: Singular values of the transformed Jacobian (see Eq. (19)).

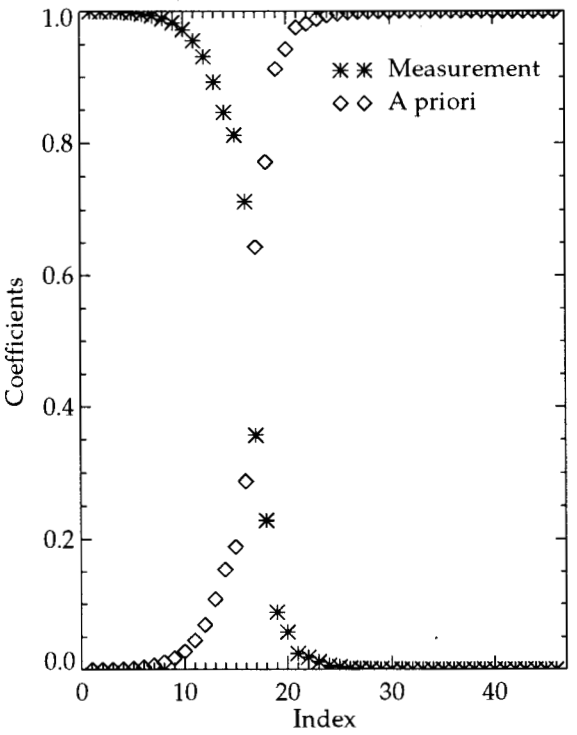

Figure 8: Proportion of the measurement and the *a priori* (see Eq. (22)).

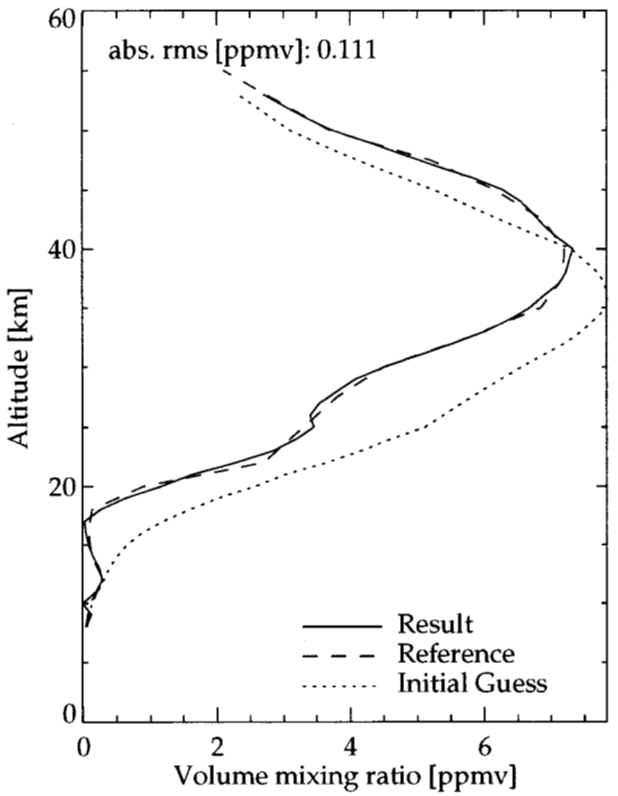

Figure 9: Result of the optimal estimation retrieval in case of an ozone-hole reference profile and a regularization constant with height.

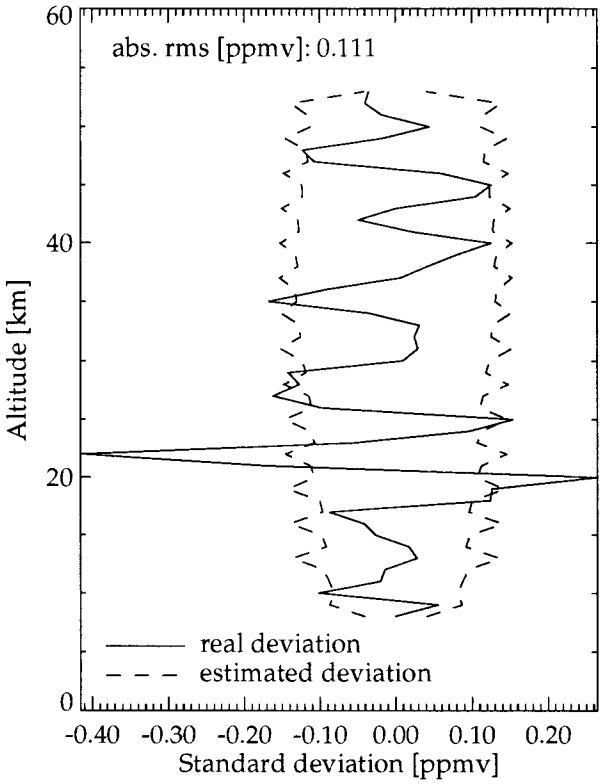

Figure 10: Real and estimated absolute deviation.

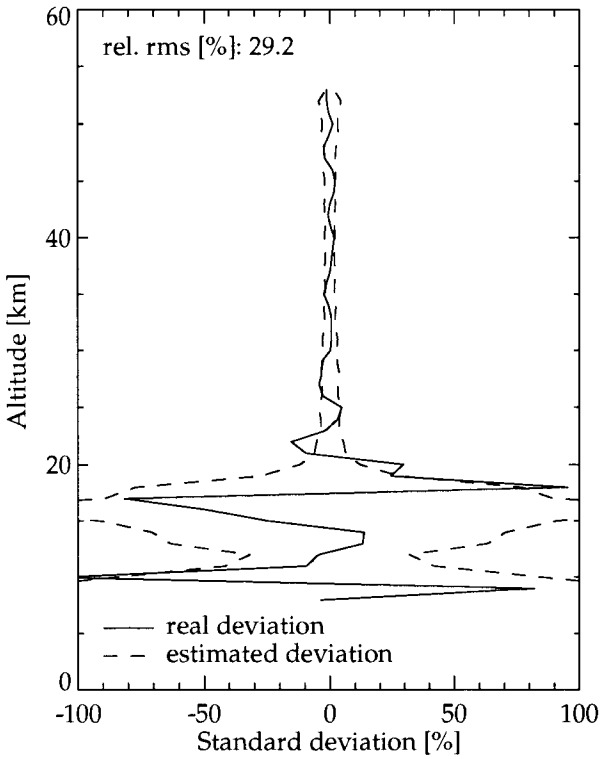

Figure 11: Real and estimated relative deviation.

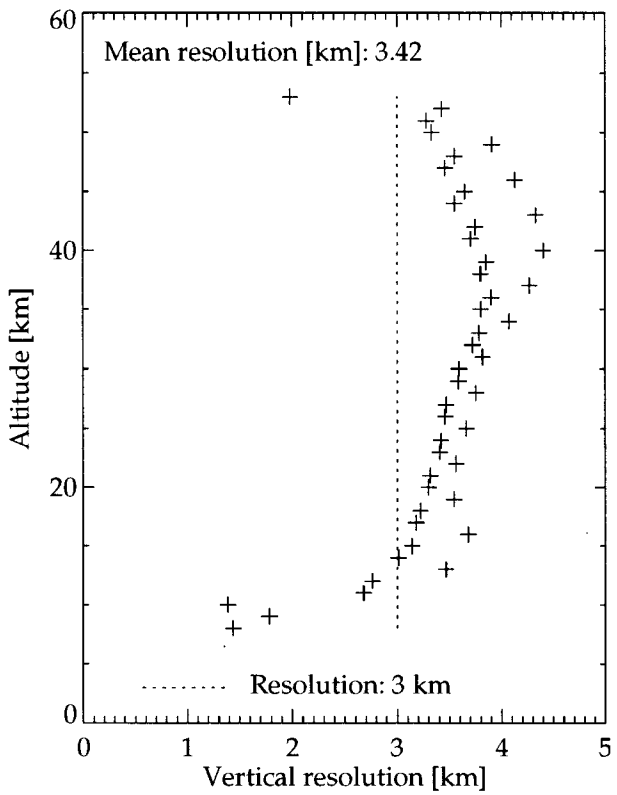

Figure 12: Vertical resolution computed via the widths of the columns of the averaging kernel.

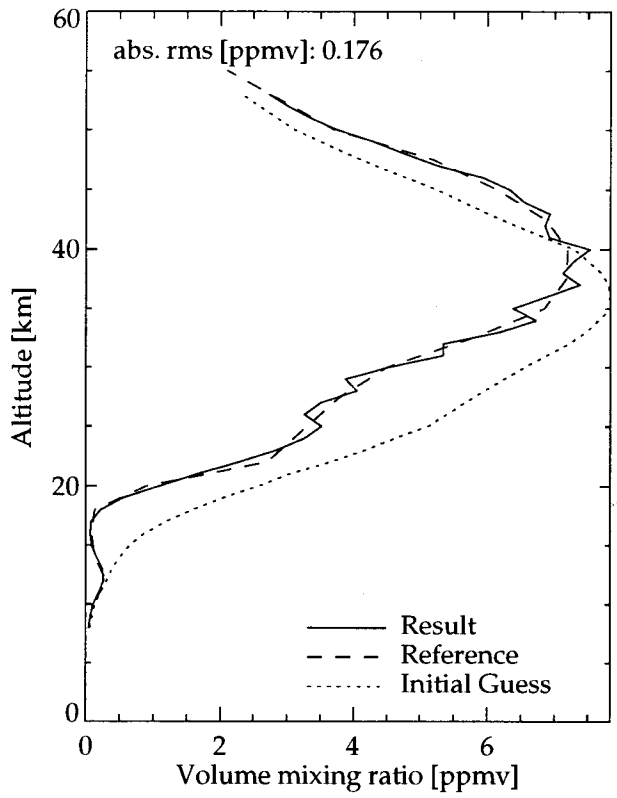

Figure 13: Result of the optimal estimation retrieval in case of an ozone-hole reference profile and a regularization dependent on height.

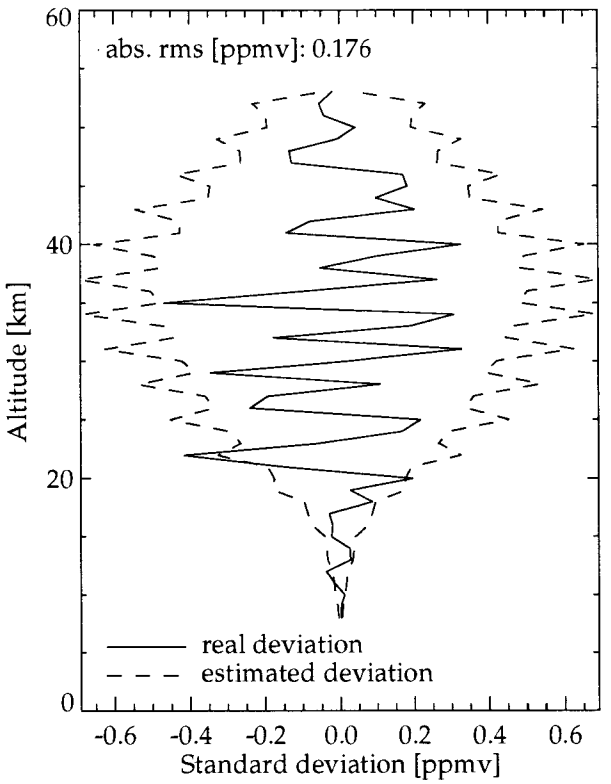

Figure 14: Real and estimated absolute deviation.

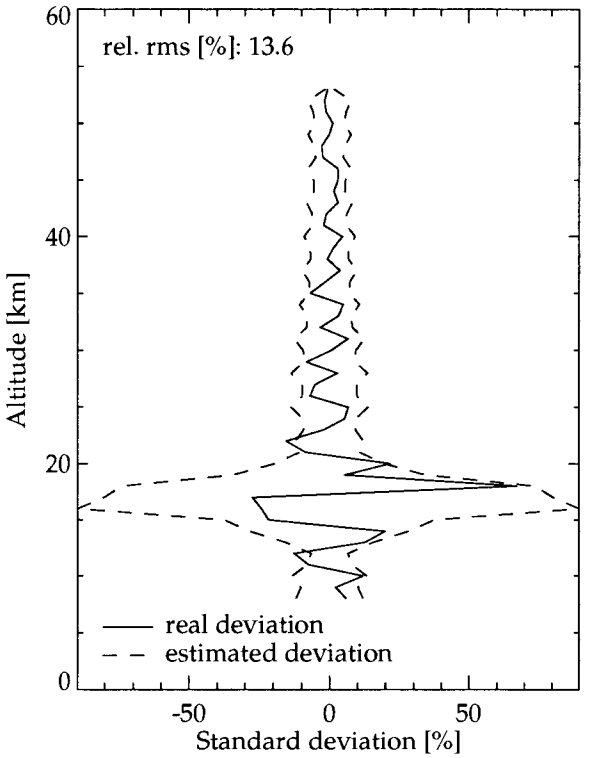

Figure 15: Real and estimated relative deviation.

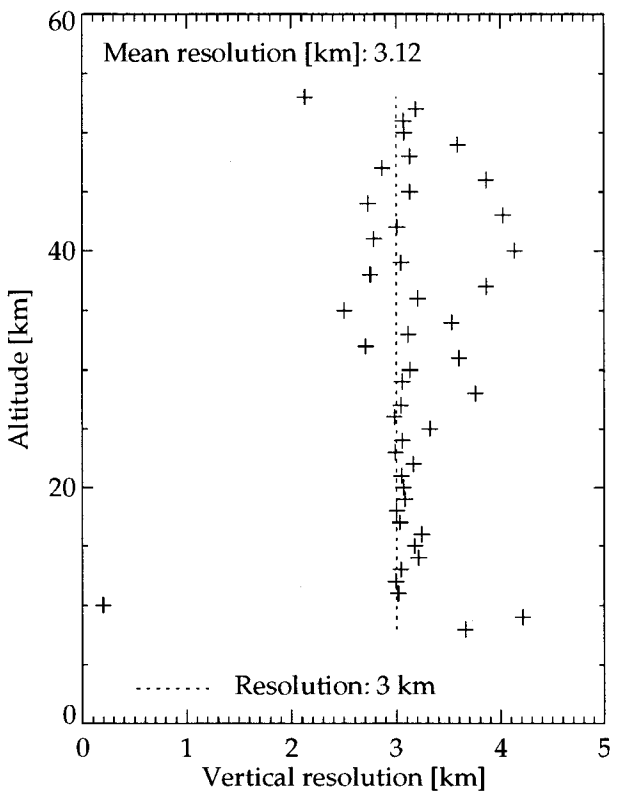

Figure 16: Vertical resolution computed via the widths of the columns of the averaging kernel.